

STUDENT EVALUATION OF TEACHING EFFECTIVENESS: AN INTRODUCTION

Kevin Brewer

Orsett Technical Reports, Series A, No.1

ISBN: 978-0-9540761-4-6

PUBLISHED BY
Orsett Psychological Services,
PO Box 179,
Grays,
Essex
RM16 3EW
UK

COPYRIGHT
Kevin Brewer 2002

COPYRIGHT NOTICE

All rights reserved. Apart from any use for the purposes of research or private study, or criticism or review, this publication may not be reproduced, stored or transmitted in any form or by any means, without prior permission in writing of the publishers. In the case of reprographic reproduction only in accordance with the terms of the licences issued by the Copyright Licensing Agency in the UK, or in accordance with the terms of licences issued by the appropriate organization outside the UK.

General Series Introduction

Orsett Technical Reports are designed to allow the exploration of specific topics in detail. Series A contains four reports on different aspects of the student evaluation of teaching effectiveness (SETE) or students' ratings of instruction (SRI). This is the rating of lecturers and teachers by their students.

REPORT No.1 ¹

This report is a literature review of the studies into SETE and SRI, mostly from the USA. The aim is to outline what students see as the "ideal lecturer". Much of the material comes from the prolific work of Kenneth Feldman.

REPORT No.2 ²

This report addresses the issue of the accuracy of students' ratings of their instructors. Is it an accurate picture of their teaching effectiveness or the personal feelings of the students? The issues of reliability, generalisability, and validity of the ratings, along with rating errors, are discussed.

REPORT No.3 ³

Report no.3 takes many of the technical issues raised in report no.2 further. In particular, the potential biases to SETE and SRI.

REPORT No.4 ⁴

This report gives details of the construction of the Birmingham Overseas Student Teaching Evaluation Questionnaire (BOSTEQ). The aim is to produce a rating instrument specifically to be used by overseas students.

The research is part of an MSc degree at the University of Aston ⁵.

¹ Brewer, K (2002a) Student evaluation of teaching effectiveness: an introduction, Orsett Technical Reports, Series A, No.1, Orsett Psychological Services: Orsett, Essex.

² Brewer, K (2002b) Student evaluation of teaching effectiveness: methodological issues - part 1, Orsett Technical Reports, Series A, No.2, Orsett Psychological Services: Orsett, Essex.

³ Brewer, K (2002c) Methodological issues with student evaluation of teaching effectiveness (SETE) - part 2, Orsett Technical Reports, Series A, No.3, Orsett Psychological Services: Orsett, Essex.

⁴ Brewer, K (2002d) Construction of Birmingham Overseas Students Teaching Evaluation Questionnaire (BOSTEQ), Orsett Technical Reports, Series A, No.4, Orsett Psychological Services: Orsett, Essex.

⁵ Brewer, K (1993) Overseas Students Evaluation of Teaching Effectiveness, Unpublished MSc thesis, University of Aston: Birmingham, UK.

CONTENTS

	Page No.
INTRODUCTION	4
IDEAL LECTURER	4
SUMMARY OF STUDIES	5
CROSS-CULTURAL STUDIES	7
DIFFERENCES AMONG FACULTIES ABOUT THE IDEAL LECTURER	8
WHAT DO STUDENTS THINK ABOUT STUDENT RATINGS OF INSTRUCTION?	11
DO STUDENTS AND TEACHERS AGREE ON WHAT IS GOOD TEACHING?	11
REFERENCES	15
APPENDIX 1 - CATEGORIES OF INSTRUCTIONAL CHARACTERISTICS USED BY FELDMAN	17
APPENDIX 2 - FACULTY DIFFERENCES FOUND IN STUDENT RATINGS OF INSTRUCTION	22
APPENDIX 3 - DETAILS OF CASHIN (1990)	24
APPENDIX 4 - DETAILS OF FELDMAN (1988)	26

INTRODUCTION

The use of student ratings of instruction seems to have started in the 1920s in North America, particularly at the University of Washington. Mueller (1951) made the first nationwide study in the USA, by approaching 1000 colleges to discover how many used student ratings: 296 were using them and planned to continue.

More recently, Cashin (1988) points out that there are over 1300 articles and books dealing with research on student ratings of teaching; the vast majority originating in North America, where student rating of instruction is common practice.

Marsh (1984) outlines the purpose of student evaluation of teaching as fourfold:

a) As a "diagnostic feedback to faculty about the effectiveness of their teaching" (p707). This is the main use for student ratings of instruction, and the issue here revolves around whether students' judgments are a true assessment of effective teaching.

b) For use in teacher promotion decisions. A controversial question in North American universities and colleges - should teaching be used as the only indicator for promotion purposes?

c) As information for students in selection of courses. Marsh and Roche (1993) feel that the Student Educational Evaluation Questionnaire (SEEQ) is not useful for inferring course quality. "Research indicates that the ratings are primarily a function of the instructor who teaches the course rather than the course" (p244).

d) To show the "outcome on a process description for research or teaching" (Marsh 1984 p707).

IDEAL LECTURER

In the first British published study in recent years, Cooper and Foy (1967) compiled a list of 43 statements relating to how a lecturer may present and organise material, their general personality, and the relationship with the class. Asking undergraduate students from all three years of a Pharmacy course, the most important characteristics chosen were "presents his material clearly and logically", "enables the student to understand the basic principles of the subject", and "can be clearly heard" in that order. Least important were

"has a democratic approach", and "is skilful in drawing blackboard diagrams."

In conclusion, the authors say that " the students placed a far higher priority than the staff on the adequate coverage of ground in the lecture course, on attempts to link lecture material to laboratory/practical work" (p185) as the main requirements of an ideal lecturer.

SUMMARY OF STUDIES

Research studies have asked students about their ideal lecturer in three main ways -

- i) To describe their ideal lecturer.
- ii) To suggest what characteristics are important for good teaching.
- iii) To describe the best teacher they have had.

Feldman (1976b) in a very extensive study records 19 important characteristics from 72 studies in North America. Depending on the type of study, slightly different characteristics are ranked as most important.

For non-structured studies (where the students volunteer the characteristics), "respect for students" is ranked highest when the students are thinking of characteristics important to good teaching; but "clarity" is first, when they are asked to think of the ideal lecturer, and "availability" when asked to think of the characteristics of the best teacher ever had.

However, in structured responses studies (where the characteristics are given), "knowledge of subject", "stimulation of interest", and "concern for class progress" are ranked highest in each type of study.

Table 1 summarises the most important characteristics based on the different types of study.

Taking the average of the ranks in the six different types of study, the most important characteristic is "knowledge of subject", followed by "stimulation of interest", and "clarity of presentation".

Feldman's work is with North American studies only. Unfortunately, there are few British studies, and certainly in recent years (since the 1970's) there is a lack of published work. Whether it is because the issue is seen as resolved, and no further research is needed, is unclear.

NON-STRUCTURED STUDIES

Characteristic Ideal Imp Best

Respect	6.5	1	2
Knowledge	5	2	3.5
Availability	-	7	1
Discussion	2	4	7
Clarity	1	10.5	5
Enthusiasm	3.5	9	11
Fairness	9.5	6	9.5
Organisation	11	12	6
Elocution	12.5	5	12
Interesting	6.5	3	3.5

STRUCTURED STUDIES

Characteristic Ideal Imp Best

Knowledge	2.5	3	1
Interesting	1	4	2.5
Class progress	4	1	5
Enthusiasm	2.5	6.5	6.5
Organisation	6	5	11
Challenge	7	8	8.5
Availability	10	10	2.5
Discussion	8	12	14
Respect	11	13	10
Clarity	5	2	6.5

KEY

Non-Structured response studies = students choose characteristics.

Structured response studies = students choose from list given.

Ideal = thinking of ideal lecturer.

Imp = thinking of characteristics important to good teaching.

Best = characteristics of best teacher ever had.

(Characteristics in full in the Appendix 1)

(Based on Feldman 1976b Table III p253)

Table 1 - Characteristics most associated with superior teachers and superior teaching, as determined by various procedures.

Furthermore, there are also hardly any British studies on the use of student evaluation of teaching effectiveness. Table 2 shows the five most important characteristics in ranked order of the four British structured studies of the ideal lecturer. The characteristics chosen are similar to the North American studies.

The only real difference is "stimulation of interest" in the North American studies, replaced by "stimulates students to think independently" in the British studies.

Thus it would seem that presentation skills and knowledge are the most desired characteristics of the ideal lecturer for students generally.

1. Clear and understandable explanations.
2. Presents material in well-organised way.
3. Shows thorough knowledge of subject.
4. Can be clearly heard in lectures.
5. Stimulates students to think independently.

Items must appear in all four studies. Based on Cooper and Foy (1967); Foy (1969); Ramsden (1975); Smithers (1970a).

Table 2 - The most important characteristics of the ideal lecturer in British studies.

CROSS-CULTURAL STUDIES

So much of the research on the evaluation of teaching has been North American, and as with many areas in psychology, cross-cultural studies are few. But in the last few years, attempts have been made to test certain rating instruments around the world.

In particular, Marsh has compared the Student Educational Evaluation Questionnaire (SEEQ), and the Endeavour Instrument in Australia (Marsh 1981; 1986), New Zealand (Watkins, Marsh and Young 1987), and Spain (Marsh, Tournon and Wheeler 1985).

Elsewhere, the same two rating instruments have been tested in Papua New Guinea (Clarkson 1984), Nepal (Watkins and Regmi 1992), Philippines (Watkins and Gerong 1992), India (Watkins and Thomas 1991), and Nigeria (Watkins and Akande 1992).

Almost all the studies found that the rating instruments clearly distinguished between the "best" and "worst" lecturers, and that the sixty-three items were seen as appropriate by at least 80% of the students.

But there were slight differences in the findings. Papua New Guinean students tended to select qualities related to the individual, rather than the organisational skills of the lecturer. In the Philippines, "good" lecturers were seen as younger and teaching major subjects, while "poor" lecturers were teaching lower grade courses and larger classes.

From New Zealand, Clift et al (1989) detail the preparation of the Teaching Performance Profile to use for promotion purposes among lecturers. The authors include separate scales for lectures, course, tutoring, lab supervision, and seminars. The items tended to load on to one global factor for the lecture scale, which included "stimulated interest", and "ability to communicate" (p200).

Also from New Zealand, a study by Watkins (1990) of student ratings used as information for perspective students in the "alternative calendar". The author found the ratings to be stable, but to vary according to size of class, year of course, and grades awarded. This does not make student ratings invalid. "But it does certainly make it difficult to compare the ratings of courses and lecturers across a university" (p19).

However, studies from the Middle East have produced different results to the North American studies. Dehlavi (1987) asked 400 Iranian students to choose their ideal lecturer from 39 characteristics. Regardless of their sex, field of study or academic year, they preferred

attributes classed as personal-social (eg: "always friendly towards students") over cognitive-intellectual (eg: "has a thorough knowledge of subject matter"). The author then tries to explain the finding as reflecting "Iranian cultural values which attach great importance to obtaining 'college degree' rather than learning per se" (p145).

Safi and Miller (1986), in the first study of student evaluation in Kuwait University, found that "punctuality in starting and finishing class" was rated highest, with "feedback on assignments and tests" rated lowest. Other important items were "pleased with students' questions", and "respected and trusted students". These three most important items are not among the highest rated in British and American studies.

Finally, Harnesh-Glezer and Meyer (1991) report a study of community college students in Israel. They build upon a study by Neumann and Neumann (1981), which showed that general satisfaction with college instruction is influenced by content and organisation of lectures; satisfaction with tests and assignments; relationship between students and staff; and satisfaction with the techniques of teaching (Harnesh-Glezer and Meyer 1991 p96).

Harnesh-Glezer and Meyer (1991) use an adapted version of a questionnaire developed by Chermesh (1978). This approach prefers a path-structure, where different aspects of course, teaching and subject are combined to give a general satisfaction score. Harnesh-Glezer and Meyer found that the major factor in course satisfaction was the teacher, with some influence from interpersonal relations among the students.

However, the authors have a word of caution: "Our results show that interpersonal relations between students are related to course satisfaction. This does not necessarily mean that they cause this satisfaction" (p105).

DIFFERENCES AMONG FACULTIES ABOUT THE IDEAL LECTURER

Two English studies, Ramsden (1975) and Smithers (1970a) break down the ratings of characteristics of the ideal lecturer into faculty groups. Though there are similarities, there are also significant differences.

Smithers divides his subject areas into Engineering, Physical Sciences, Life Sciences, and Arts/Social Sciences. They all agree that "knows his subject" and "presents his subject so that students can understand it"

are most important, while "keeps strictly to the text-books" is least desirable.

But Physical Sciences, and Arts/Social Sciences rate significantly lower than the other two faculty groups, the following: "speaks slowly enough for full notes to be taken" and "appears to enjoy lecturing". Life Sciences, and Arts/Social Sciences rate significantly higher, "concentrates on providing the essentials of the subject as a framework for independent study".

Finally, Engineering students rate significantly higher, "appears to like students" than do Physical Sciences and Arts/Social Sciences, and significantly lower than them, "takes his own distinctive line on a controversial matter or where a problem has yet to be resolved."

Smithers summarises the difference as "students of applied science and engineering appeared to look towards lecturers for information, social scientists, for stimulation." (p141)

Ramsden compares 3 faculty groups: (1) science, engineering and mathematics; (2) arts and social science; and (3) professional courses, at what was North East London Polytechnic. He notes three significant differences in their ratings:

i) Science, engineering and mathematics students compared to arts and social science students rate higher on "adjusts his or her pace to the needs of the class" and "is clear and understandable in his or her explanations", but lower on "is sensitive to feelings and problems of individual students".

ii) Students on professional courses rate lower "adjusts pace" and "sensitive to feelings" and higher on "shows the relevance of his or her subject to the work you expect to do when you qualify" and "gives a good factual coverage of the subject matter" than science, engineering and maths students.

iii) Arts and social science students rate lower "shows relevance" than students on professional courses.

From the North American studies, Riley et al (1950) found for the arts faculty - "knowledge of subject" and "ability to encourage thought" are most important; for Biological and Physical Sciences - "ability to explain" and "organisation"; and for the Social Sciences - "ability to encourage thought" and "organisation".

While Birney, Coplin and Grose (1960) similarly found slight differences between faculty. Students of the Humanities see "dynamic and stimulating" and "helpful to students" as the most important; students of Natural

Science prefer "organised" and "helpful"; Social Sciences faculty suggest "dynamic" and "knowledge of subject" (quoted in Feldman and Newcomb 1969a). The full results of both these studies are in Appendix 2.

Cashin (1990) comparing responses to the Student Instructional Report (SIR), and the Instructional Development and Effectiveness Assessment (IDEA) produces 44 different academic fields from 87 843 classes in 316 colleges and universities in North America.

Taking the average ratings on each department, the academic fields were divided into five categories, from high to low. The high rating departments tend to be arts and humanities, except English Literature and Language, and History, which are medium low. The low raters are mostly business, economics, computer science, maths, physical science and engineering. The full details are in Appendix 3.

In a more recent, American, study Jennings and Bartling (1991) found significant differences between courses in the following areas: more maths and natural science majors responded "most definitely" to "A professor should try to encourage joint student/faculty research projects" than business majors.

Business majors preferred courses related to professional experiences, while natural and social science majors wanted courses related to research.

Cashin proposes six reasons for the difference in ratings by faculty:

i) Quantitative courses receive lower ratings - students on these courses have less developed verbal skills, and thus these courses are more difficult to teach.

ii) Sequential courses (courses following each other in terms of building on knowledge eg science) receive lower ratings - students have a less solid foundation of knowledge.

iii) How masterable is course content? - certain courses are difficult to define in terms of their content (eg: American History), and thus allow greater freedom, leading to higher ratings.

iv) Academic fields vary in academic skills and goals assessed as important.

v) Purpose of the course - difficult courses aimed at weeding out weak students may be rated lower.

vi) Some academic fields are poorly taught.

"Probably the real explanation lies in some combination of the explanations just offered" (Cashin 1990 p119).

Because of faculty differences in ratings, Barnes and Barnes (1993) suggest the same evaluation instrument may not be applicable.

Erdle and Murray (1986) took a different approach, looking at teachers from different faculties to see if their teaching behaviour varied. Trained observers watched 124 teachers.

"Interpersonal orientation" behaviour occurred more frequently with Arts and Social Science teachers than Natural Science teachers, while "task oriented" behaviour occurred more with Natural Science and Social Science teachers than Arts teachers. However, for students, what makes an effective teacher was the same across faculties.

WHAT DO STUDENTS' THINK ABOUT STUDENT RATINGS OF INSTRUCTION?

Costin et al (1971) used a questionnaire with 404 psychology students at the University of Illinois. The findings were generally favourable - only 14% agreed that "ratings are a waste of time"; 49% were willing to spend more time on rating classes; and 64% wanted to rate more classes. The full results are in table 3.

Overall, then, students feel positive about their rating of instruction. But faculty are not so convinced.

DO STUDENTS AND TEACHERS AGREE ON WHAT IS GOOD TEACHING?

"Students strongly believed that student evaluations do measure teaching effectiveness and that they were qualified to evaluate their teachers, while faculty only slightly agreed with these statements" (McBean and Al-Nassri 1982 p278).

Getting students to evaluate their classes has overlooked the question of whether the staff regard the same things as important as the students. The easiest way to discover this information is to ask students and staff to rate the same lecturer, and then find the correlation between the two groups. The higher the positive correlation between the students' and staff ratings, the more they agree.

Items	% OF RESPONSES				
	Strong agree	Some agree	Neutral	Some disagree	Strong disagree
1.The forms will affect most teacher's future performance	3	42	14	24	11
2. The forms will affect most teacher's departmental status or advancement	2	12	12	31	28
3. Ratings are a waste of time	4	10	9	42	33
4. I would like to be able to rate more classes	32	32	25	5	5
5. I would be willing to spend time outside of class to rate courses	17	32	16	13	17
6. In retrospect, I would rate a number of courses differently now than at the time I took them	6	27	13	33	16
7. My ratings usually agree with those of the rest of the class	1	23	16	12	2
8. I generally rate a professor higher than he deserves, since there are so few good professors	2	10	10	29	46
9. A professor who tells good jokes will get a good rating	3	24	15	30	25

(Based on Costin et al 1971 p 523)

Table 3 - students' opinions concerning student rating of instruction (N=404).

Ramsden (1975) asked 11 members of staff along with their classes to rank 23 statements for importance with the ideal lecturer. A significant Spearman rank correlation coefficient of 0.715 was found between student and staff rankings. But the key differences were for staff: most stress on being "well prepared for classes" and "stimulating students to think independently" - and for students: more emphasis on "ability of lecturer to stress important material".

Feldman (1988) in a detailed literature review looked at 31 studies from North America. The average correlation between faculty and students was $r = .71$. However, this correlation of the general similarity between faculty and students disguises differences. Looking at 22 individual characteristics of teaching, Feldman found differences on rating of importance, of seven. Table 4 shows the top ten ranked characteristics

for students and staff. Students rank "sensitivity" as number one (number three for staff), followed by "preparation" (number four for staff). "Knowledge of subject" was most important for staff, and "enthusiasm" next. More details of this study appear in Appendix 4.

STUDENTS	FACULTY
1. Teacher's sensitivity to, concern with, class level and progress	1. Teacher's knowledge of subject
2. Teacher's preparation: organisation of course	2. Teacher's enthusiasm (for subject or for teaching)
3. = Teacher's knowledge of subject	3. Teacher's sensitivity to, and concern with, class level and progress
3. = Teacher's stimulation of interest in course and its subject matter	4. Teacher's preparation: organisation of the course
5. Teacher's enthusiasm (for subject or for teaching)	5. Clarity and understandableness
6. Clarity and understandableness	6. Intellectual challenge and encouraging of individual thought (by teacher and course)
7. Teacher's availability and helpfulness	7. = Instructor's fairness; impartiality of evaluation of students; quality of exam
8. Teacher's concern and respect for students; friendliness of teacher	7. = Teacher motivates students to do their best; high standard of performance required
9. Perceived outcome or impact of instruction	9. = Teacher's availability and helpfulness
10. Instructor's fairness; impartiality of evaluation of students; quality of examination	9. = Teacher's concern and respect for students; friendliness of teacher

(Based on Feldman 1988 p311)

Table 4 - Importance (rank order) of various instructional dimensions for students and faculty.

A comprehensive recent report is by Williams, Loder and Fry (1993), which questioned 2357 students, 2309 academics, and 1222 administrators in the UK. It was a general survey of life in higher education, but one question asked what characteristics contribute to effective teaching. The results are detailed in table 5, which shows the slight differences in rank order of

characteristics between students, faculty and administrators. In this case, students rate "knowledge" highest, while the academics and administrators prefer "enthusiastic about subject". Large differences in ranking appear on "challenge students' thinking" (ranked seven for students, but ranked two for academics); and "detail and prompt feedback"(ranked three for students, and only six for academics).

Generally there is agreement between staff and students, and certainly agreement about which teachers are good, but the disagreement is over why they are rated as good (Hildebrand and Wilson 1970).

Characteristics	Students	Academics	Administrators
Knowledge of subject	1	3	2
Enthusiastic about subject	2	1	1
Detailed and prompt feedback	3	6	3
Good presentational skills	4	5	4
Consciously transmit knowledge to students	5	7	7
Encourage active participation by students	6	4	5
Challenge students' thinking	7	2	6
Put knowledge of Learning Theory into practice	8	-	9
Provide reading lists to guide student reading	9	-	-
Set students regular assessed work	-	8	8

Ranked items are those rated "very important".

(Based on Williams et al 1993 Table 1).

Table 5 - Characteristics of lecturers contributing to effective teaching as assessed by three sources.

REFERENCES

- Barnes, L.L.B & Barnes, M.W(1993) Academic discipline and generalizability of student evaluations of instruction, *Research in Higher Education*, 34, 2, 135-149
- Cashin, W.E(1988) *Student Ratings of Teaching: A Summary of the Research*, Kansas State University: Center for Faculty Evaluation and Development
- Cashin, W.E(1990) Students do rate different academic fields differently, *New Directions for Teaching and Learning*, 43, Fall, 113-121
- Chermesh, R(1978) Instructional evaluation by students: its use for diagnosis and improvement of teaching, *Educational Technology*, 18, 9-13
- Clarkson, P.C(1984) Papua New Guinean students' perceptions of mathematics lecturers, *Journal of Educational Psychology*, 76, 6, 1386-1395
- Clift, J; Hall, C & Turner, I(1989) Establishing the validity of a set of summative teaching performance scales, *Assessment and Evaluation in Higher Education*, 14, 3, 193-206
- Cooper, B & Foy, J.M(1967) Evaluating the effectiveness of lectures, *Universities Quarterly*, 21, 2, 182-185
- Costin, F; Greenough, W.T & Menges, R.T(1971) Student ratings of college teaching: reliability, validity and usefulness, *Review of Educational Research*, 41, 511-535
- Dehlavi, N.S(1971) "Iranian students' perception of ideal teacher, *Perceptual and Motor Skills*, 64, 143-146
- Erdle, S & Murray, H.G(1986) Interfaculty differences in classroom teaching behaviours and their relationship to student instructional ratings, *Research in Higher Education*, 24, 2, 115-127
- Feldman, K.A(1976b) The superior college teacher from the students' view, *Research in Higher Education*, 5, 243-288
- Feldman, K.A(1988) Effective college teaching from the students' and faculty's view: matched or mismatched priorities? *Research in Higher Education*, 28, 4, 291-344
- Feldman, K.A(1989b) The association between student ratings of specific instructional dimensions and student achievement: refining and extending the synthesis of data from multi-section validity studies, *Research in Higher Education*, 30, 6, 583-646
- Feldman, K.A(1993) College students' views of male and female college teachers Part 1 - evidence from the social laboratory and experiments, *Research in Higher Education*, 33, 3, 317-375
- Feldman, K.A & Newcomb, T.M(1969a) *Impact of College on Students Vol I*, San Francisco: Jossey-Bass
- Foy, J.M(1969) A note on lecturer evaluation by students, *Universities Quarterly*, 23, 3, 345-349
- Harnesh-Glazer, M & Meyer, J(1991) Dimensions of satisfaction with college education, *Assessment and Evaluation in Higher Education*, 16, 2, 95-107
- Jennings, C.M & Bartling, C.A(1991) Quality education: what college students think *College, Student Journal*, 25, 1, 532-536
- McBean, E.A & Al-Nassri, S(1982) Questionnaire design for student measurement of teaching effectiveness, *Higher Education*, 11, 273-288
- Marsh, H.W(1981) Students' evaluations of tertiary

instruction: testing the applicability of American surveys in an Australian setting, *Australian Journal of Education*, 25, 2, 177-193

Marsh, H.W.(1984) Students' evaluation of university teaching: dimensionality, reliability, validity, potential biases and utility, *Journal of Educational Psychology*, 76, 5, 707-754

Marsh, H.W.(1986) Applicability paradigm: students' evaluations of teaching effectiveness in different countries, *Journal of Educational Psychology*, 78, 6, 465-473

Marsh, H.W; Tournon, J & Wheeler, B(1985) Students' evaluations of university instructors: the applicability of American instruments in a Spanish setting, *Teaching and Teacher Education*, 1, 2, 123-138

Mueller, F.J(1951) Trends in student ratings of faculty, *American Association of University Professors' Bulletin*, 37, 319-324

Neumann, Y & Neumann, E(1981) Determinants of student satisfaction with course work: an international comparison between two universities, *Research in Higher Education*, 14, 321-333

Ramsden, P(1975) Polytechnic students' expectations of their teachers and the use of a student feedback questionnaire: a preliminary report, *Higher Education Bulletin*, 3, 2, 73-85

Riley, J.W; Ryan, B.F & Lifschitz, M(1950) *The Student Looks At His Teacher*, New Brunswick, NJ: Rutgers University Press

Safi, Q & Miller, R.I(1986) Student evaluations of courses and instructors at Kuwait University, *Higher Education Review*, 18, 3, 17-26

Smithers, A(1970a) Some factors in lecturing, *Educational Review*, 22, 2, 141-150

Watkins, D(1990) Student ratings of tertiary courses for "alternative calendar" purposes, *Assessment and Evaluation in Higher Education*, 15, 1, 12-21

Watkins, D & Akande, A(1992) Student evaluations of teaching effectiveness: a Nigerian investigation, *Higher Education*, 24, 453-463

Watkins, D & Gerong, A(1992) Evaluating undergraduate college teaching: a Filipino investigation, *Educational and Psychological Measurement*, 52, 3, 727-734

Watkins, D & Regmi, M(1992) Student evaluations of tertiary teaching: a Nepalese investigation, *Educational Psychology*, 12, 2, 131-142

Watkins, D & Thomas, B(1991) Assessing teaching effectiveness: an Indian perspective, *Assessment and Evaluation in Higher Education*, 16, 3, 185-198

Watkins, D; Marsh, H.W & Young, D(1987) Evaluating tertiary teaching: a New Zealand perspective, *Teaching and Teacher Education*, 3, 1, 41-53

Williams, G; Loder, C & Fry, H(1993) *Identifying priorities*, Centre for Higher Education Studies Newsletter, 3; London: Institute of Education

APPENDIX 1

CATEGORIES OF INSTRUCTIONAL CHARACTERISTICS AS USED BY FELDMAN

For each instructional dimension, examples of evaluative items that would be classified into it are given.

1. Stimulation of interest:

- the instructor puts material across in an interesting way;
- the instructor gets students interested in the subject;
- it was easy to remain attentive;
- the teacher stimulated intellectual curiosity.

2. Instructor's Enthusiasm:

- the instructor shows interest and enthusiasm in the subject;
- the instructor seems to enjoy teaching;
- the teacher communicates a genuine desire to teach students;
- the instructor never showed boredom for teaching this class;
- the instructor showed energy and excitement.

3. Knowledge of subject matter:

- the instructor has a good command of the subject material;
- the teacher has a thorough knowledge, basic and current, of the subject;
- the instructor has good knowledge about or beyond the textbook;
- the instructor knows the answer to students' questions;
- the teacher keeps lecture material updated.

4. Intellectual expansiveness:

- the teacher is well informed in all related fields;
- the teacher has respect for other subject areas and indicates their relationship to his or her own subject of presentation;
- the teacher exhibited a high degree of cultural attainment.

5. Preparation/organisation:

- the teacher was well prepared for each day's lecture;
- the presentation of the material is well organised;
- the overall development of the course has good continuity;
- the instructor planned the activities of each class period in detail.

6. Clarity:

- the instructor gave clear explanations;
- the instructor interprets abstract ideas and theories clearly;
- the instructor makes good use of examples and illustrations to get across difficult points;
- the teacher effectively synthesizes and summarises the material;
- the teacher answers students' questions in a way that helps students to understand.

7. Elocution:

- the instructor has a good vocal delivery;
- the teacher speaks distinctly, fluently and without hesitation;
- the teacher varied the tone of his or her voice;
- the teacher has the ability to speak distinctly and be clearly heard;
- the instructor changed pitch, volume and quality of speech.

8. Sensitivity:

- the teacher was skilled in observing students' reaction;
- the teacher was aware when students failed to keep up in class;
- the instructor teaches near the class level;
- the teacher takes an active personal interest in the progress of the class and shows a desire for students to learn.

9. Course Objectives:

- the purposes and policies of the course were made clear to the student;
- the instructor gave a clear idea of student requirements;
- the teacher clearly defined student responsibilities in the course;
- the teacher tells students which topics are most important and what they can expect on tests;
- the instructor gave clear assignments.

10. Course Materials:

- the teacher has the ability to apply material to real life;
- the instructor makes the course practical;
- there is worthwhile and informative material in lectures that doesn't duplicate the text;
- the course has excellent content;
- the class considers what we are learning worth learning.

11. Supplementary Materials:

- the homework assignments and supplementary readings were helpful in understanding the course;
- the teacher made good use of teaching aids such as films and other audio-visual materials;
- the instructor provided a variety of activities in class and used a variety of media (slides, films, projections, drawings) and outside resource persons.

12. Outcome:

- gaining of new knowledge was facilitated by the instructor;
- I developed significant skills in the field;
- I developed increased sensitivity and evaluative judgment;
- the instructor has given me tools for attacking problems;
- the course has increased my general knowledge;
- apart from your personal feelings about the teacher, has he/she been instrumental in increasing knowledge of the course's subject matter.

13. Fairness:

- grading in the course was fair;
- the instructor has definite standards and is impartial in grading;
- the exams reflect material emphasized in the course;
- test questions were clear;
- coverage of subject matter on exams was comprehensive.

14. Personality:

- the teacher has a good sense of humour;
- the teacher was sincere and honest;
- the teacher is highly personable at all times in dress, voice, social grace, and manners;
- the instructor was free of personal peculiarities;
- the instructor is not autocratic and does not try to force us to accept his ideas and interpretations;
- the teacher exhibits a causal, informal attitude;
- the instructor laughed at his own mistakes.

15. Feedback:

- the teacher gave satisfactory feedback on graded material;
- criticism of papers was helpful to students;
- the teacher told students when they had done a good job;
- the teacher is prompt in returning tests and assignments.

16. Encouragement of Questions:

- students felt free to ask questions or express opinions;
- the instructor stimulated class discussion;
- the teacher encouraged the students to express differences of opinions and to evaluate each other's ideas; the instructor invited criticism of his or her own ideas;
- the teacher appeared receptive to new ideas and the viewpoints of others.

17. Intellectual Challenge:

- this course challenged students intellectually;
- the teacher encouraged students to think out answers and follow up ideas;
- the teacher attempted to stimulate creativity;
- the instructor raised challenging questions and problems.

18. Respect:

- the instructor seems to have a genuine interest in and concern for students;
- the teacher took students seriously;
- the instructor established good rapport with students;
- the teacher was friendly toward all students.

19. Availability:

- the instructor was willing to help students having difficulty;

- the instructor was willing to give individual attention;
- the teacher was available for consultation;
- the teacher was accessible to students outside class.

20. Motivated students:

- instructor motivates students to do their best work;
- the instructor sets high standards of achievement for students;
- the teacher raises the aspirational level of students.

21. Self-Initiated Learning:

- students are encouraged to work independently;
- students assume much responsibility for their learning;
- the general approach used in the course gives emphasis to learning on the students' own;
- the teacher does not suppress individual initiative.

22. Research Productivity:

- the talks about his own research;
- instructor displays high research accomplishments;
- the instructor publishes material related to his subject field.

23. Difficulty (description):

- the workload and pace of the course was difficult;
- I spent a great many hours studying for this course;
- the amount of work required for this course was very heavy;
- this course required a lot of time;
- the instructor assigned very difficult reading.

24. Difficulty (evaluation):

- the content of this course is too hard;
- the teacher's lectures and oral presentations are over my head;
- the instructor often asked for more than students could get done;
- the instructor attempted to cover too much material and presented it too rapidly.

25. Classroom Management:

- the instructor controls class discussion to prevent rambling and confusion;
- the instructor maintained a classroom atmosphere conducive to learning;
- students are allowed to participate in deciding the course content;
- the teacher did not rule with an iron hand.

26. Classroom Atmosphere:

- the class does not make me nervous;
- I felt comfortable in this class;
- the instructor created an atmosphere in which students in the

- class seemed friendly;
- this was not one of those classes where students failed to laugh, joke, smile, or show other signs of humour;
- the teacher is always criticising and arguing with students.

27. Individualization:

- instead of expecting every student to do the same thing, the instructor provides different activities for different students;
- my grade depends primarily on my improvement over my past performance;
- in this class each student is accepted on his or her own merits;
- my grade is influenced by what is best for me as a person as well as by how much I have learned;
- the instructor evaluated each student as an individual.

28. Pursuit of Objectives:

- the instructor accomplished what he or she set out to do;
- there was close agreement between the announced objectives of the course and what was actually taught;
- course objectives stated agreed with those actually pursued.

Based on Feldman (1989b; Appendix A pp633-636). Original details in Feldman (1976b): only 19 characteristics; number 12 was "difficulty of course, and number 14 "classroom management".

Feldman (1993) adds:

- 29. Overall Rating of Lectures as an Item of a Multi-item Indicator.
- 30. Overall Rating of Teacher as an Item of a Multi-item Indicator.
- 31. Overall Rating of Course as an Item of a Multi-item Indicator.

APPENDIX 2

FACULTY DIFFERENCES FOUND IN STUDENT RATINGS OF INSTRUCTION

% of students checking as one of the three
most important qualities to good teaching.

Qualities Thought to be Important to Good Teaching	Teachers of: Arts	Social Sciences	Biological & Physical Sciences
Ability to encourage thought	47	17	70
Expert knowledge of subject	54	70	42
Systematic organisation of subject matter	32	78	48
Ability to explain clearly	42	89	38
Enthusiastic attitude toward subject	46	16	26
Good speaking ability	31	6	11
Tolerance toward student disagreement	11	2	45
Fairness in marking and grading tests	8	10	7
Pleasing personality	16	4	5
Sympathetic attitude toward students	9	5	3

(Based on Feldman and Newcomb 1969a p256)

Table showing qualities chosen as one of the three most important
qualities to good teaching based on faculty from Riley et al
(1950).

Rank of frequency of mention

Characteristics of Best Teachers	Teachers of:		
	Humanities	Natural Science	Social Science
Dynamic, stimulating	1	4	1
Interested in and helpful to students	2	2	3
Clear, well organised	5.5	1	6
Skilled technique, well prepared	7	3	4
Holds interest	3	7	5
Broad knowledge, thorough mastery (of field)	5.5	9	2
Great interest in and devotion to field	4	8	8
Fine personality	8	6	7
Makes connection to other subjects	9	5	9

(Based on Feldman and Newcomb 1969a p257)

Table showing characteristics of best teacher for each of three subject matters according to students at Amherst, from Birney, Coplin and Grose (1960).

APPENDIX 3

Level of Rating	Course Effect	Instructor Effect
HIGH	Architecture; Art; Communications; Fine and Applied Arts; Health and Technology; Home Economics; Music; Secretarial Studies; Speech	Area Studies; Fine and Applied Arts; Foreign Languages; Military Science: Music
MEDIUM HIGH	Agriculture; Education; Foreign Languages; Letters and Humanities; Library Science; Nursing; Physical and Health Education; Political Science and Government; Public Affairs; Religion and Theology; Social Work and Service; Trade and Vocational Technology Education	Art; Communications; English Language and Literature; Health and Technology; History; Home Economics; Law; Letters and Humanities; Library Science; Physical and Health Education; Political Science and Government; Psychology; Public Affairs; Religion and Theology; Secretarial Studies; Speech; Trade and Vocational Technology Education
MEDIUM	Area Studies; Biological Sciences; Health Professions; Interdisciplinary Studies; Military Science; Psychology	Business and Commercial Technology (*); Education (*); Engineering Technology: Nursing (*)
MEDIUM LOW	Accounting; Business and Commercial Technology; English Language and Literature; History; Law; Social Science; Sociology	Agriculture; Biological Science; Chemistry; Health Professions; Maths and Statistics; Philosophy; Social Science; Social Work and Service; Sociology

Level of Rating	Course Effects	Instructor Effects
LOW	Business and Management; Chemistry; Computer and Information Science; Data-processing Technology; Economics; Engineering; Engineering Technology; Maths and Statistics; Philosophy; Physical Science; Physics	Accounting; Architecture; Business and Management; Computer and Information Science; Data-processing Technology; Economics; Engineering; Interdisciplinary Studies; Physical Science: Physics

(*) = received mixed rankings, so average

(Based on Cashin 1990; table 1 pp116-7)

Table showing a summary of ratings on course effectiveness and instructor effectiveness for IDEA and SIR academic-field data.

APPENDIX 4

Feldman (1988) finds similarities between faculty and student ratings in the particular teaching characteristics of "enthusiasm"; "knowledge"; "expansive"; and "preparation". But students rate higher "stimulation".

SIMILARITIES IN RATINGS BETWEEN STUDENTS AND FACULTY

Characteristic	Number of studies	Import stated by students (a)	Import stated by faculty (b)	Difference (b-a)
1.Stimulation	10	.28	.50	+.22
2.Enthusiasm	9	.32	.24	-.08
3.Knowledge	17	.28	.23	-.05
4.Expansive	9	.56	.50	-.06
5.Preparation	12	.27	.28	+.01
6.Clarity	8	.33	.39	+.06
7.Elocution	9	.47	.57	+.10
8.Sensitivity	7	.22	.27	+.05
9.Objectives	10	.63	.56	-.07
10.Materials	11	.46	.47	+.01
11.Materials	10	.54	.58	+.04
12.Outcome	4	.43	.51	+.08
13.Fairness	14	.45	.45	-
14.Personality	12	.64	.70	+.06
15.Feedback	9	.49	.53	+.04
16.Questions	15	.51	.48	-.03
17.Challenge	10	.58	.40	-.18
18.Respect	16	.39	.47	+.08
19.Availability	10	.37	.47	+.10
20.Motivates	7	.58	.45	-.13
21.Self-learn	4	.75	.51	-.24
22.Research	5	.91	.88	-.03

Standardised ranks eg important in 20 of 37 studies = 0.5; smaller = more important. Difference: + = students more important. Characteristics in detail in Appendix 1.

Table showing the similarity between students and faculty on individual characteristics by standardised ranks found by Feldman (1988).